

基于 SA-BPNN 多模态融合的教学质量评价方法

王文静^{1,2}, 范涛^{1,2*}, 王国中^{1,2}, 赵海武^{1,2}

(1. 上海工程技术大学电气工程学院, 上海 201600; 2. 人工智能产业研究院, 上海 201600)

摘要: 针对传统评估教学未充分利用教学过程信息, 导致泛化能力差和预测精度低等问题, 提出一种基于 SA-BPNN 多模态融合的教学质量评价模型。该模型包括模态特征提取模块和多模态融合预测模块。首先构建全过程多维度的教学质量评价体系, 包括人工评分、在线教育平台和教学视频 3 种模态数据, 通过自注意力机制改进反向传播神经网络模块(SA-BPNN)提取各模态特征信息。其次鉴于评教过程存在模态数据缺失的可能, 在预测模块的早期融合阶段引入多模态混合融合策略以改进反向传播神经网络(MF-BPNN), 该策略按照特定规则将不同模态的特征信息混合再融合, 以减轻模型对某一模态数据的依赖。随后将融合后的特征信息输入到 MF-BPNN 模块中, 生成最终的评教结果。该模型融合多模态信息进行评价, 避免了传统评教的主观性, 且混合融合策略提升了模态缺失时的预测精度。在某高校真实数据集上的实验结果表明, 相较于 BPNN 和 GA-BPNN 方法, 该模型的均方误差分别提高了 2.4~3.9 个百分点, 能够高效准确地评估教学质量, 为检验教师教学效果和优化教学管理, 最终全面提升教学质量提供理论支持。

关键词: 质量评价; 自注意力机制; 多模态融合; 特征提取; 神经网络

中图分类号: TP391

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0068448

Teaching Quality Evaluation Method Based on SA-BPNN Multi-Modal Fusion

WANG Wenjing^{1,2}, FAN Tao^{1,2*}, WANG Guozhong^{1,2}, ZHAO Haiwu^{1,2}

(1. School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201600, China;

2. Artificial Intelligence Industry Research Institute, Shanghai 201600, China)

【Abstract】 A Simulated Annealing Backpropagation Neural Network (SA-BPNN)-based multi-modal fusion model for evaluating teaching quality is proposed to address the issues of poor generalization ability and low prediction accuracy caused by the insufficient utilization of teaching process information in traditional teaching evaluations. The model comprises two parts: a modal feature extraction module and a multi-modal fusion prediction module. First, a multidimensional teaching quality evaluation system with three modal data types: manual scoring, online education platforms, and teaching videos, is constructed. The SA-BPNN module is improved using a self-attention mechanism to extract feature information from each modality. Second, considering the possibility of missing modal data in the evaluation process, a multi-modal hybrid fusion strategy is introduced in the early fusion stage of the prediction module to improve the MF-BPNN. This strategy combines and fuses the feature information of different modalities according to specific rules to reduce the reliance of the model on a certain modality. The fused feature information is then input into the MF-BPNN module to generate the final evaluation results. The model that fuses multi-modal information for evaluation avoids the subjectivity of traditional evaluations, and the hybrid fusion strategy improves the prediction accuracy when modal data are missing. Experimental results based on the real dataset of a university show that, compared with the BPNN and GA-BPNN methods, the Mean Square Error (MSE) of this model increases by 2.4 and 3.9 percentage point, respectively, which can effectively and accurately evaluate the teaching quality and provide theoretical support for testing the teaching effect of teachers, optimizing teaching management, and comprehensively improving the teaching quality.

【Key words】 quality evaluation; self-attention mechanism; multi-modal fusion; feature extraction; neural network

0 引言

教学质量是学校生存与繁荣的核心, 也是评估教学管理和教学成果的重要方面。教学质量评价是利用教育评价理论和技术对教学过程及其结

果是否达到一定质量要求做出价值判断^[1]。通过对教师教学活动的定量评价, 可以客观地评估教学效果, 检验教学目标是否达到预期^[2], 有助于教育管理机构有效监督教学质量, 从而全面提升高校教育水平。教学质量评价涉及的内容广泛、指

收稿日期: 2023-09-25 修回日期: 2023-11-23

基金项目: 国家重点研发计划(2019YFB180270200)。

通信作者 E-mail: *02200001@sues.edu.cn

标繁多,是模糊且非线性的问题。评价时常出现评价指标体系不完善、评价方法不科学等问题^[3]。因此,探索更加智能高效的教学质量评价方法具有重要意义。

国外教学评估的研究起源于 20 世纪 20 年代。1927 年,黑麦斯首次发表了高等学校教学评估的系列研究报告,深入探讨了教学评估的效度和信度问题。随后在 20 世纪 30 年代,发达国家开始对课程教学质量进行评价,并涌现出各种评价模式。众多学者在借鉴专家评教经验的基础上,构建了教学质量评价体系和评价方法,这为高等教育机构管理者提升教学质量水平提供了有力的工具^[4]。

评价指标与评价方法相互交织,共同决定了教学质量评价结果的准确性。评价指标的制定标准关乎是否能够准确评估教育质量,从而深刻影响受教育者的培养品质。评价指标体系被定义为对被评价对象的质量和数量方面进行具体评价的集合,其本质上是依据评价目的而制定的评价内容的综合体^[5]。目前,相关学者从多个不同视角探讨,并没有形成明确一致的标准。文献^[6]结合教学现状以及评价原则,提出评价指标应包括教师教学评价、学生学习评价、学习资源评价以及教学条件 4 个方面。文献^[5]运用结构方程模型和层次分析法得出评价指标体系应涵盖教学态度、教学内容、教学方法、教学规范、教学组织、教学能力和教学效果等 7 个重要方面。文献^[7]从课前准备、课中实施、课后效果 3 个维度,根据教学过程确定了教学评价指标内容。同时,文献^[8]还探讨了在互联网时代,如何从教师、学生和教学资源等多重角度出发,精细化地构建教学质量评价体系,以应对当前教育模式的挑战。除了前述的典型研究,还有研究主要聚焦于特定类型的高校^[9-10]、特定学科领域^[11]或特定教学情境^[12-13]。然而,这些评价体系侧重于评估教学要素,只考虑单一要素的质量标准,忽略整体教学质量的特性^[14],体系的构建方法缺乏科学性和系统性,因此评价结果的有效性较低^[15-16]。

为了适应传统的主观定性评估场景,大多数相关研究采用模糊综合评价、层次分析法和神经网络等方法建立带有权重的教学质量评估模型。文献^[17]采用层次分析法来确定指标的权重,并建立一种基于马尔可夫链的模糊动态综合评估模型。文献^[18]将基于表达式的 AHP 群决策模型应用于教学质量评价,并采用有序加权平均算子

将多位专家在多个指标上的评价集结为综合评价。文献^[19]采用多元统计方法构建了评价本科课堂教学质量的指标体系,并运用群决策的层次分析法确定指标的权重以计算最终评价结果。上述方法虽充分体现了专家经验知识,但评估过程存在随机性和主观性,未能考虑到评价指标与教学效果之间的非线性复杂关系,导致评价结果未能真实反映教学质量。

本文运用现代信息技术分析和挖掘了教学场景和环节的评价数据,建立了多模态教学质量评价体系,并提出了基于 SA-BPNN 多模态融合的评价模型。该模型由自注意力机制改进的 BP 神经网络(SA-BPNN)和多模态融合策略改进的 BP 神经网络(MF-BPNN)两个模块组成。其中,SA-BPNN 模块可以捕捉同一模态不同评价指标之间的关联性,为每个指标分配动态权重从而突出重要信息,有效增强指标间特征学习和关系建模能力。MF-BPNN 模块基于混合融合的算法思想,经过加权的表示作为高级特征输入,在输入层将各模态的高级特征信息进行混合重分组,组内进行加权融合,摆脱了在实际应用场景下模型对模态信息完整性的依赖。

本文的主要贡献如下:

- 1)设计一种多模态特征提取模块。该模块基于自注意力机制改进 BP 神经网络,可以有效针对评教数据的特点,增强教学质量评价指标间的特征学习能力,对数据进行预训练,得到各模态的特征信息。
- 2)提出多模态特征融和策略。该策略将预处理后的模态特征信息进行混合再融合,改善局部模态缺失情况下模型的应变能力,增强模型的鲁棒性。
- 3)在某高校真实数据集上进行实验,与近期具有代表性的教学质量评价方法相比取得更好的实验结果。

1 相关研究

1.1 基于神经网络的教学质量评价方法

新一代人工智能技术(如目标检测、自然语言处理、姿态识别等)结合大数据和互联网+等新技术,正在引领数据驱动的教学质量评价方法研究^[20-21]。神经网络作为信息科学与技术研究领域的一种计算模型,模拟人脑结构和功能,实现了自学习和非线性变换等功能,从而弥补了先前方法的不足。

文献^[22]采用 BP 神经网络的自学习和结构

特点,开发了教学质量评价模型,解决主观因素和评价复杂性问题。文献[23]提出了基于 GA-BPNN 算法的英语教学质量评价模型,通过引入遗传算法优化 BP 神经网络,从而提升预测精度。文献[3]使用改进粒子群优化(IPSO)算法优化了 BP 神经网络预测模型,且构建了以教学态度、内容、方法、效果为指标的评价体系。文献[24]提出一种基于层次分析法和粒子群优化 BP 神经网络的在线教学质量评价模型,利用层次分析法确定各指标的权重,通过粒子群优化算法优化的 BP 神经网络建立模型,根据各指标的权重系数计算评分结果。文献[25]研究了在线教学质量评价机器学习方法,引入自适应学习速率和动量项,优化 BP 神经网络的下降算法,显著提高了模型的收敛速度,同时提出一种新的深度神经网络模型,有效地处理了复杂的高维大规模数据集。

然而,上述研究受制于单一模态主观量化的指标,大多在传统 BP 神经网络基础上进行一定的算法优化,未充分考虑实际的教学方法与教学场景,在与人工智能技术的结合上仍有很大的探索空间,在实际应用过程中具有一定的局限性,难以在实际教学工作中提供动态精准的指导建议。

1.2 多模态特征提取与融合

对于包含复杂语义信息的场景,多模态数据可以更好地表征场景的特征,从而为后续的模型决策提供有效信息。随着教育信息化进程的不断加快,除了基本的人工评教数据和在线教育平台数据,各种课堂观察技术也为教学质量评价提供了大量数据。如文献[26]提出的基于改进 OpenPose 的学习者行为识别方法,文献[27]提出的基于深度图像的多学习者姿态识别方法,文献[28]提出的基于多模态数据的课堂学生行为识别技术的探究等。同时,不同模态的评教数据存在差异,衡量的侧重点也存在一定程度的不同。针对不同的模态分别构建模型进行特征提取,强化模型对特定模态的特征提取能力,有助于模型决策能力的提升。

根据任务、数据类型和数据来源的不同,多模态特征融合技术分为早期融合、晚期融合和混合融合^[29]。现有的教学质量评价模型通常采用晚期融合方法,传统的晚期融合常在决策阶段使用加权融合来计算各模态的贡献权重,从而确定最终结果。然而,加权融合方法存在两个问题:首先,难以准确量化各模态对结果的贡献以及确定相应的权重,可能导致某一模态特征对预测结果的影响过大;其次,这种方法不能进行端到端模型训练,对于局部模态

缺失时的应对能力有限。

而混合融合即可用于跨模态的信息融合,又可以改善部分模态缺失情况下的预测效果,为端到端的训练模型提供可能。文献[30]采用混合融合方法实现了一种多媒体事件检测算法,该方法结合早期与晚期融合,在捕捉模态关系的同时减轻了过拟合问题,显著提升了模型的性能。组合策略的设计与合理性是混合融合方法的关键问题,将神经网络的灵活性与高容错性运用于混合融合方法,具备深入研究的价值。

目前主流教学质量评价研究流程如图 1 所示。

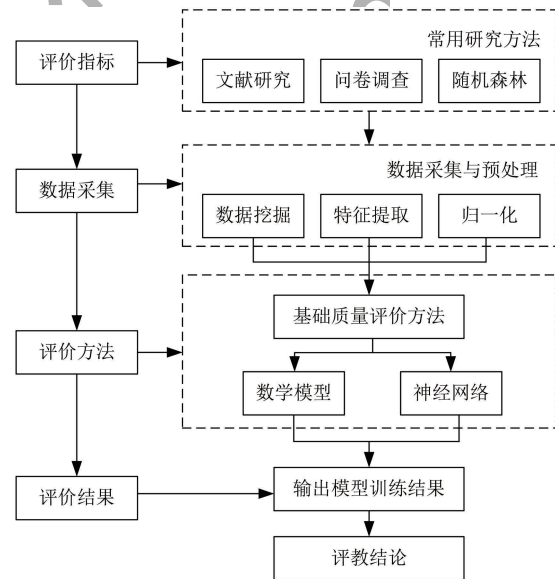


图 1 教学质量评价研究流程

Fig. 1 Procedure of educational quality evaluation research

2 模型设计

本文面向人工评分、在线教育平台和课堂视频 3 种评教数据模态(以下简称人工、平台和视频),采用层级模型结构分别获取各模态信息特征后进行混合融合,实现了多模态融合评价。本文提出一种基于自注意力机制改进 BP 神经网络的多模态融合教学质量评价模型。该模型包含自注意力机制、模态特征提取、多模态特征融合以及评价预测等 4 个组成部分。模型结构如图 2 所示。

2.1 数据预处理

2.1.1 人工模态

针对人工模态数据,本文基于上海工程科技大学的评教问卷设定采集指标,数据由督导组、教师或学生进行随堂评价收集,人工模态评价指标体系如表 1 所示。表 1 中的人工模态 E_1 展开的 8 项评价指标,即 $e_{11} \sim e_{18}$,作为数据集采集项,将一组人工评价数据表示为 $x_i = [e_{11}, e_{12}, \dots, e_{1i}]$, $i = 8$ 。

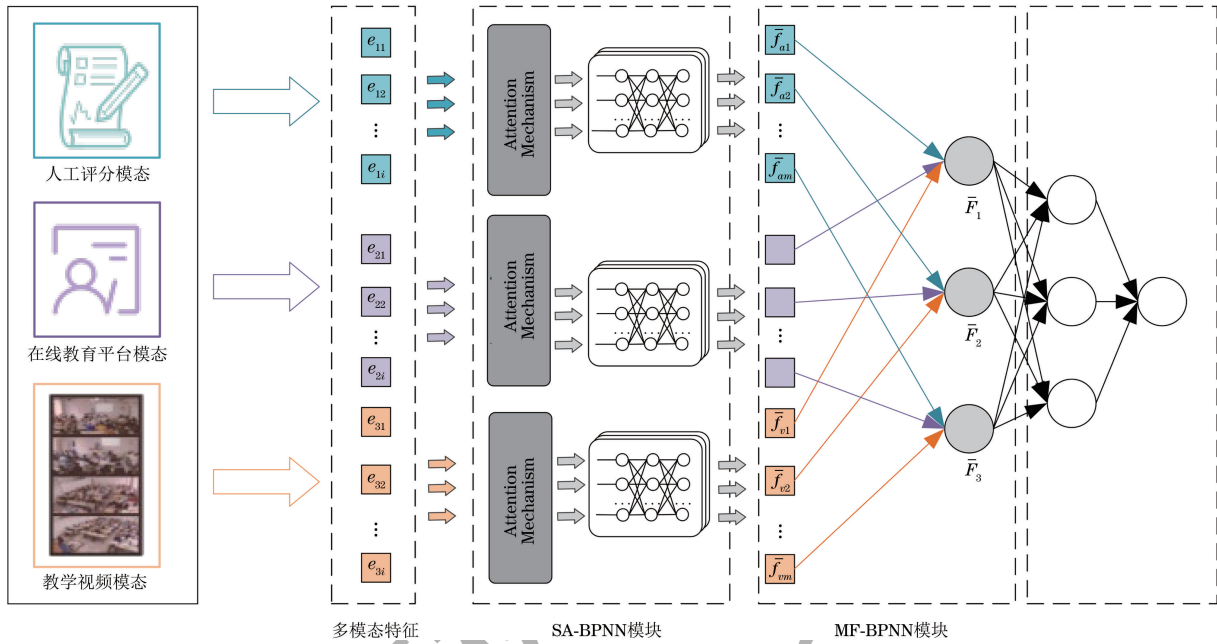


图 2 本文模型结构

Fig. 2 Structure of the model in this paper

表 1 人工模态评价指标体系

Table 1 Manual modality evaluation index system

评价模态	教学阶段	评价指标
人工评分	教学设计	教学内容、教学工具准备充分
		教学目标明确、重点突出 教学方法合理、内容充实
	教学过程	理论与实际相结合
		课堂气氛活跃、学生表现积极 重视课堂纪律、课堂活动有序
		教学形式多样、创新性强
	教学效果	依据教学大纲达成教学目标

2.1.2 平台模态

针对平台模态数据,本文以北京世纪超星学习通、科大讯飞畅言智慧课堂以及上海工程技术大学爱课堂^[31]这 3 款普遍具备在线学习功能的软件进行大数据分析,平台模态评价指标体系如表 2 所示。表 2 中的在线教育平台模态 E_2 展开的 9 项评价指

表 2 平台模态评价指标体系

Table 2 Platform modality evaluation index system

评价模态	教学阶段	评价指标
在线教育平台	教学设计	教学资源更新频次
		发布课程资源数量
	教学过程	师生互动频次
		迟到、早退频次
		发起考勤频次
		布置作业频次
		作业完成率
	教学效果	作业批改率
		访问教学资源频次

标,即 $e_{21} \sim e_{29}$,作为数据集采集项,将一组平台评价数据表示为 $x_j = [e_{21}, e_{22}, \dots, e_{2j}]$, $j = 9$ 。

2.1.3 视频模态

针对视频模态数据,本文使用自建教学监控视频数据集采集评价数据,自建视频模态数据集如图 3 所示,视频模态评价指标体系如表 3 所示。该数据集是从某高校选取的真实课堂监控提取,出于隐私安全协议的考虑,在本文中所展示的数据集图片均进行了模糊处理。表 3 中的视频模态 E_3 展开的 7 项评价指标,即 $e_{31} \sim e_{37}$,作为数据集采集项。将一组平台评价数据表示为 $x_k = [e_{31}, e_{32}, \dots, e_{3k}]$, $k = 7$ 。

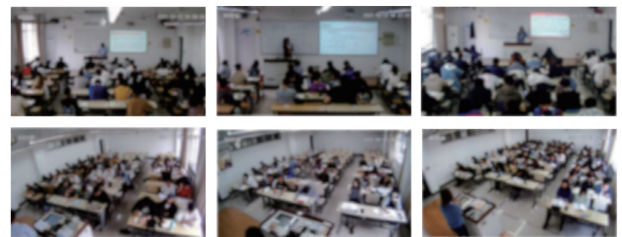


图 3 自建教学视频模态数据集示例

Fig. 3 Example of self-constructed teaching video modality dataset

表 3 视频模态评价指标体系

Table 3 Video modality evaluation index system

评价模态	教学阶段	评价指标
课堂视频	教学过程	课堂互动参与人数
		板书时长
		迟到、早退人数
	教学效果	课堂出勤时长
		出勤率
		抬头率
		教室前三排入座率

2.1.4 数据归一化

由于不同指标之间的评价单位不同,本文使用最大最小值法对数据进行归一化处理,将归一化后的模态数据表示为 $X_i = [e_{11}, e_{12}, \dots, e_{1r}]$ 、 $X_j = [e_{11}, e_{12}, \dots, e_{1j}]$ 和 $X_k = [e_{11}, e_{12}, \dots, e_{1k}]$, 其中, X_i 表示人工模态输入数据, X_j 表示平台模态输入数据, X_k 表示视频模态输入数据。

2.2 模态特征提取

SA-BPNN 模块旨在捕捉 3 种不同模态的特征信息。引入自注意力机制的目的在于增强教学质量评价体系中的特征学习和关系建模能力。自注意力机制在有效捕捉不同评价指标之间关联性的同时,为每个指标分配动态权重,突出重要信息,并适应长距离依赖关系。这些经过加权的表示作为高级特征重新输入,从而提升模型性能和预测准确性。

自注意力机制核心思想在于“非局部平均”,以提升重要信息的关注度和特征之间的整体联系。其核心是通过计算目标位置 Q 与其他非局部位置 K 的信息相似性来信息传递。加权相加在非局部区域内进行,从而有效修正目标位置的信息。该机制分为 3 个步骤,每个特征存在 Q 、 K 和 V 3 个矩阵。

首先,以模态 a 为例,分别是 $Q_i = WX_i^a W_1$ 、 $K_i = WX_i^a W_2$ 、 $V_i = WX_i^a W_3$ 。计算特征之间的相似性如式(1)所示:

$$S_j = \frac{Q_i \times K_j}{\sqrt{d_k}} \quad (1)$$

式中: X_i^a 表示 a 模态的第 i 个特征信息; Q_i 表示第 i 个特征信息的 Q 矩阵; K_i 表示第 i 个特征的 K 矩阵; $\sqrt{d_k}$ 表示对相似度进行缩放的值,确保相似度计算在适当范围内; S_j 表示 i 第个特征信息与第 j 个特征信息间的相似度; W 、 W_1 、 W_2 和 W_3 为引入的可训练参数,用于对模型进行训练以适应特定任务。

其次,应用 Softmax 函数对相似性进行归一化处理得到权重表示。这些权重表示相似度在整体输入中的权重分布,总和为 1。权重值 Z_i 如式(2)所示:

$$Z_i = \frac{\exp S_j}{\sum_{j=1}^m \exp S_j} \quad (2)$$

式中:变量 m 表示 a 模态的特征数量; \exp 表示以自然常数 e 为底的指数函数。

最后,计算模态 a 的第 i 个特征与其他特征之间的注意力值 f_{a_i} ,如式(3)所示:

$$f_{a_i} = \sum_{j=1}^m Z_j V_j \quad (3)$$

同理,3 种模态的高级特征信息最终计算结果如式(4)所示:

$$\begin{cases} F_a = [f_{a_1}, f_{a_2}, f_{a_3}, \dots, f_{a_n}] \\ F_p = [f_{p_1}, f_{p_2}, f_{p_3}, \dots, f_{p_n}] \\ F_v = [f_{v_1}, f_{v_2}, f_{v_3}, \dots, f_{v_n}] \end{cases} \quad (4)$$

将这些高级特征信息作为输入传递到后续的 BP 神经网络中,具体如式(5)所示:

$$\bar{f}_j = g\left(\sum_{i=1}^m W_{ij} f_{a_i}\right) \quad (5)$$

式中: W_{ij} 表示权重矩阵; f_{a_i} 表示模态 a 的第 i 个高级特征信息;函数 $g(\cdot)$ 为激活函数; \bar{f}_j 表示第 j 个神经元的信息。

最终,所获得的人工模态、平台模态和视频模态特征信息 \bar{f} 表示如式(6)所示:

$$\begin{cases} \bar{F}_a = [\bar{f}_{a_1}, \bar{f}_{a_2}, \bar{f}_{a_3}, \dots, \bar{f}_{a_m}] \\ \bar{F}_p = [\bar{f}_{p_1}, \bar{f}_{p_2}, \bar{f}_{p_3}, \dots, \bar{f}_{p_m}] \\ \bar{F}_v = [\bar{f}_{v_1}, \bar{f}_{v_2}, \bar{f}_{v_3}, \dots, \bar{f}_{v_m}] \end{cases} \quad (6)$$

式中: m 表示模态特征信息的维数。

2.3 特征融合策略

MF-BPNN 模块是一种可训练的混合融合结构,用于跨模态信息融合,改进了在部分模态缺失情况下模型的预测效果。

已知本文混合融合前且当模态信息完整时,输入信息组成如式(7)所示:

$$\{F_{\text{input}}\} = \begin{cases} \bar{F}_a = [\bar{f}_{a_1}, \bar{f}_{a_2}, \bar{f}_{a_3}, \dots, \bar{f}_{a_m}] \\ \bar{F}_p = [\bar{f}_{p_1}, \bar{f}_{p_2}, \bar{f}_{p_3}, \dots, \bar{f}_{p_m}] \\ \bar{F}_v = [\bar{f}_{v_1}, \bar{f}_{v_2}, \bar{f}_{v_3}, \dots, \bar{f}_{v_m}] \end{cases} \quad (7)$$

对特征信息进行混合融合,预测模型输入层第 j 个神经元的输入信息计算如式(8)所示:

$$\{F_{\text{input}}\}_j = \frac{1}{n} \sum_{i=1}^n \bar{f}_{ij} \quad (8)$$

式中: n 表示测试样本输入模态的数量; i 表示第 i 种模态。

本文人工、平台和视频 3 种模态混合融合后的输入信息如式(9)所示:

$$\{F_{\text{input}}\} = \begin{cases} \bar{F}_1 = [\bar{f}_{a_1} + \bar{f}_{p_1} + \bar{f}_{v_1}]/3 \\ \bar{F}_2 = [\bar{f}_{a_2} + \bar{f}_{p_2} + \bar{f}_{v_2}]/3 \\ \vdots \\ \bar{F}_3 = [\bar{f}_{a_m} + \bar{f}_{p_m} + \bar{f}_{v_m}]/3 \end{cases} \quad (9)$$

2.4 教学质量评价预测

教学质量评价预测模型基于 BP 神经网络结构。第一层是包含 m 个神经元的输入层,输入表示

为 $F_{input} = \{\bar{F}_1, \bar{F}_2, \dots, \bar{F}_j\}$, 第二层是包含 p 个神经元的隐藏层, 最后一层是包含 q 个神经元的输出层, 输出值为 $y_{score} = y$ 。输入层与隐藏层神经元的计算过程如式(10)所示:

$$h_i = \varphi \left(\sum_{j=1}^m \omega_{ij} \bar{F}_j + \theta_i \right), i = 1, 2, \dots, p \quad (10)$$

式中: h_i 表示第 i 个神经元的输出值; φ 表示激活函数 Sigmoid; θ_i 表示第 i 个神经元的线性偏置; ω_{ij} 表示第 i 个神经元与第 j 个神经元的连接权值。

Tansig 函数是非线性单调函数, 常用于表示非线性关系, 同时满足反向传播神经网络中梯度下降算法的可微性要求^[32]。本文选用 Tansig 函数作为激活函数, 记为 $\varphi(x)$, 隐藏层与输出层神经元的关系如式(11)所示:

$$o_k = \varphi \left(\sum_{i=1}^p \omega_{ik} h_i + \theta_k \right), k = 1, 2, \dots, q \quad (11)$$

神经网络的训练本质是通过比较网络输出和目标输出之间的误差来学习数据之间的相关性。常用的评价指标包括均方误差(MSE)、交叉熵损失、准确率和 R^2 等。在训练阶段, 损失函数采用 MSE, 如式(12)所示:

$$MSE_j = \frac{1}{2m} \sum_{k=1}^n (y_k - o_k)^2 \quad (12)$$

式中: y_k 表示评教结果的真实值; o_k 表示评教结果的预测值; m 表示测试样本模态的数量; n 为测试样本的数量。

重复训练步骤, 最终输出对教师的教学质量评分预测值, 训练后的神经网络可以实现从输入到输出的映射。

3 实验设计

3.1 实验设置

关于 BP 神经网络的拓扑结构, 目前尚未找到较好的构造方法。输入层节点数与教学质量评价体系中的指标个数对应。为便于观察和调整模型的训练效果, 隐含层设置为 1 层, 隐含层节点数通过实验误差进行调整, 最终设定为 5。输出层节点数设置为 1, 即教学质量评价预测结果。学习率设置为 5×10^{-5} , 最大迭代次数设置为 100, 最小误差设为 0.000 1, 训练将在达到最大迭代次数或预设精度误差时停止。

本文选取上海工程技术大学 2021 年到 2022 年 20 门公开课程作为教学质量评价对象, 所选课程包括大学英语、高等数学和中国近现代史纲要等。数据集包含学生评教问卷数据、每门课程每学期的课

堂监控视频数据(分辨率为 1 080 像素)和爱课堂平台(在线教育平台)使用数据。教务处督导专家根据这 3 种模态的数据对教师教学质量进行综合评分, 将其结果作为模型的期望输出。共收集样本总量为 2 000 份, 划分为训练集(80%)、验证集(10%)和测试集(10%)。为了增强本文模型对于模态缺失情况下的准确性, 在训练阶段中, 每个模态输入的数据有 1% 的概率不参与训练。BP 神经网络模型在对特征信息进行训练时, 实验中的数据都是在同一设置下重复运行 5 次后取结果的平均值得到的。

实验系统软件环境如表 4 所示。

表 4 实验系统软件环境

Table 4 Experimental system software environment

参数类型	参数
系统	Linux 服务器
处理器	Intel Core i9-10900K CPU 和 Nvidia RTX A4000 16 GB GPU
安装内存	16 GB
应用软件版本	PyCharm V3.3

3.2 实验结果与分析

3.2.1 消融实验

为了验证本文所提出的改进方法的有效性, 本文在视频、平台、人工 3 种模态下的自建数据集上进行了 4 个不同的实验设置, 实验结果如表 5 所示。其中, MSE_AVG(10^{-3}) 表示 5 次实验的均方误差结果的平均值, MSE_MIN(10^{-3}) 表示 5 次实验中最佳的均方误差结果, 均值融合是将 3 种模态下输出的预测结果取平均值作为综合评价结果。实验结果表明, 使用 SA-BPNN 模块以及 MF-BPNN 模块均能有效地提高模型的预测精度。

表 5 消融实验结果

Table 5 Results of ablation experiments

序号	神经网络类型	融合模块	MSE_AVG	MSE_MIN
1	BPNN	—	0.95	0.71
2	BPNN	MF-BPNN	0.67	0.49
3	SA-BPNN	无缝融合	0.23	0.07
4	SA-BPNN	MF-BPNN	0.18	0.16

实验数据分析结果表明, SA-BPNN 模型在预测效果方面明显优于实验 1 和实验 2 中的方法。在实验 3 和实验 4 中, 经过 MF-BPNN 模块处理后的预测模型相对于处理前的效果更为稳定。

3.2.2 对比实验

为了验证教学质量评价指标体系的科学性以及 SA-BPNN 多模态融合模型(SAMF-BPNN)的可靠性, 本研究采用 MSE 作为评估性能指标, 在人工、

平台和视频 3 种模态类型下使用同一数据集进行实验。将本文提出的模型与经典 BP 神经网络、GA-BPNN^[23] 改进模型以及近三年内基于 BP 的最新研究改进模型,包括 PSO-BP^[24]、DA-BP^[33] 和 IPSO-BP^[34] 模型进行实验对比。当训练误差小于目标误差时训练结束,结果如表 6 所示。实验结果表明,相较于其他模型,SA-BPNN 模型具有更快的收敛速度,能够在较少的迭代次数内找到最优解。同时,该模型的训练误差较小,能够获得更好的预测准确率,实验结果如图 4、图 5 所示。

表 6 不同模型的对比实验结果

Table 6 Comparative experimental results of different models

序号	神经网络类型	MSE_AVG	MSE_MIN
1	BP	0.144	0.114
2	GA-BP	0.113	0.106
3	PSO-BP	0.130	0.115
4	DA-BP	0.146	0.131
5	IPSO-BP	0.129	0.117
6	SAMF-BPNN	0.093	0.087

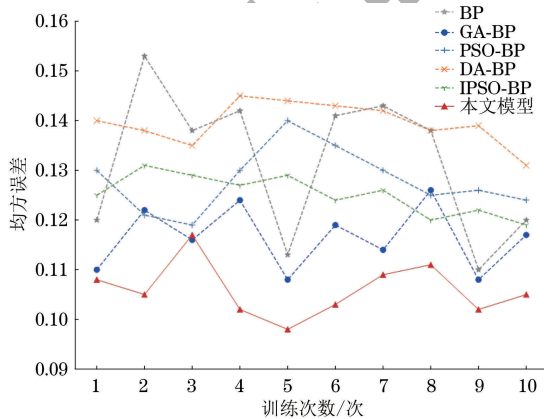


图 4 不同神经网络训练过程的均方误差

Fig. 4 Mean square error of different neural network training process

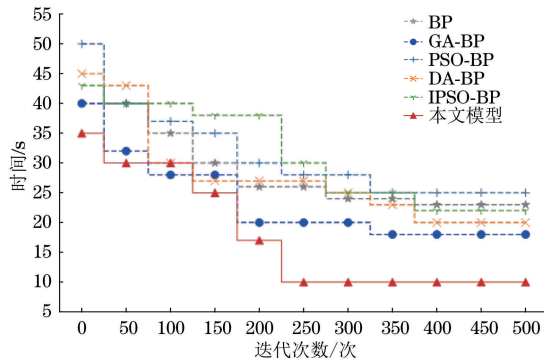


图 5 不同神经网络的运行时间

Fig. 5 Running times of different neural networks

其次,分别使用 GA-BP 神经网络、SA-BP 神经网络对测试组随机抽取 15 组数据进行预测,预测结果如图 6 所示。

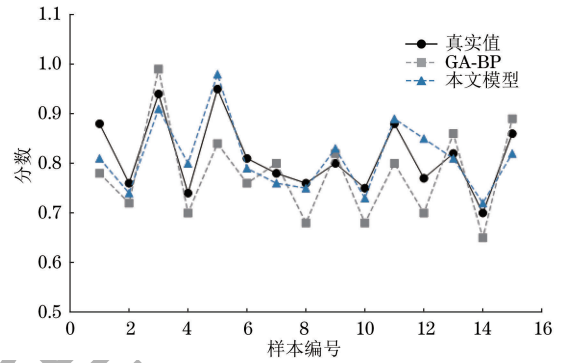


图 6 三种神经网络的预测结果

Fig. 6 Predictive results of three neural networks

从以上实验结果可以看出,基于 SA-BPNN 多模态融合的教学质量评价模型可以显著加快网络的收敛速度,并且在较少的迭代次数中找到最优解,同时获得更好的预测精度,确保实验的训练结果更稳定,从而提高教学质量评价系统在运行过程中的使用效果。

3.2.3 混合融合模块验证

本节验证在部分模态缺失时 MF-BPNN 模块的实用性,以测试集某一样本课程为例,实验结果如表 7 所示。

表 7 部分模态缺失时的实验结果

Table 7 Experimental results with partial modality missing

序号	模态类型	真实值	SA-BP+混合融合	SA-BP
1	视频	0.95	0.691 6	0.342 1
2	平台	0.95	0.784 5	0.573 2
3	人工	0.95	0.915 3	0.825 7
4	视频/人工	0.95	0.929 1	0.524 3
5	视频/平台	0.95	0.874 3	0.489 7
6	人工/平台	0.95	0.937 6	0.861 8
7	视频/人工/平台	0.95	0.946 6	0.965 6

实验结果表明,本文提出的 MF-BPNN 模块可以在有限的代价下允许部分模态的缺失,增强了教学质量评价模型的鲁棒性。综上,可以认为基于 SA-BPNN 多模态融合的教学质量评价方法在教学质量评价场景中具有良好的应用价值。

4 结束语

本文针对传统教学质量评价存在的问题,提出一种基于 SA-BPNN 多模态融合的教学质量评价方法。该方法基于 SA-BPNN 模块和 MF-BPNN 模块,引入可训练的混合融合策略,为实现端到端训练的教学质量评价预测模型提供了更大的可能。实验结果表明,相较于传统的 BP 神经网络和 GA-BP 神经网络,所提出的模型在预测结果上精度有所提高。

本文所设计的 MF-BPNN 模块可以提高实际教学场景中部分模态数据缺失时模型预测的准确性,为后续教学质量评价系统的设计提供更灵活多样的方案选择。下一步研究可以探索更多的模态数据融合方法,进一步提高评价模型的准确性和可靠性。此外,可以将该模型应用于更广泛的教育领域,例如在线教育和远程学习评价等,以验证其在不同教学场景下的适用性和效果。

参考文献

- [1] 陈玉琨. 教育评价学[M]. 北京: 人民教育出版社, 1999.
CHEN Y K. Educational evaluation[M]. Beijing: People's Education Press, 1999. (in Chinese)
- [2] YANG J, SHAO Z. Fragility and countermeasure analysis of teaching quality evaluation mechanism in ethnic region[J]. Journal of Research on Education for Ethnic Minorities, 2021, 32(5): 150-158.
- [3] CHEN L Y, WANG L H, ZHANG C Y. Teaching quality evaluation of animal science specialty based on IPSO-BP neural network model[J]. Computational Intelligence and Neuroscience, 2022, 2022: 3138885.
- [4] 郑尚男. 基于卷积神经网络的高校线上教学质量评价模型及应用研究[D]. 大庆: 东北石油大学, 2023.
ZHENG S N. Research on the evaluation model and application of online teaching quality in colleges and universities based on convolutional neural network [D]. Daqing: Northeast Petroleum University, 2023. (in Chinese)
- [5] 徐向君. 高校院系专业结构调整与认证及教学质量保障指导手册下[M]. 北京: 中国教育出版社, 2008.
XU X J. Handbook on the adjustment of university departmental and major structures, accreditation, and guidance on teaching quality assurance, volume II [M]. Beijing: China Education Publishing House, 2008. (in Chinese)
- [6] 朱兴荣. 试析网络空间课堂教学评价指标体系的构建[J]. 电脑知识与技术, 2015, 11(22): 109-111.
ZHU X R. On the construction of evaluation index system for classroom teaching in cyberspace[J]. Computer Knowledge and Technology, 2015, 11(22): 109-111. (in Chinese)
- [7] 薛宁, 杨晓冬. 高职院校线上教学质量评价指标体系的构建与实践[J]. 高等职业教育探索, 2021, 20(2): 75-80.
XUE N, YANG X D. Construction and practice of online teaching quality evaluation index system in higher vocational colleges [J]. Higher Vocational Education Exploration, 2021, 20(2): 75-80. (in Chinese)
- [8] 李倩. 互联网时代背景下线上线下混合式教学质量评价体系构建——以地方应用型本科高校为例[J]. 中国新通信, 2022, 24(13): 179-181.
LI Q. Construction of online and offline mixed teaching quality evaluation system under the background of Internet era: a case study of local application-oriented undergraduate universities [J]. China New Telecommunications, 2022, 24(13): 179-181. (in Chinese)
- [9] 刘自团, 彭华安. “双一流”建设高校教学质量现状及影响因素研究——基于江苏省 15 所高校的调查与分析[J]. 中国高教研究, 2020(8): 23-29.
LIU Z T, PENG H A. An study on the current situation and influencing factors of teaching quality in “double first-class” construction universities: investigation and analysis on 15 universities in Jiangsu Province[J]. China Higher Education Research, 2020(8): 23-29. (in Chinese)
- [10] 张炜, 邓勇新, 辛越优, 等. 多元分类视角的高等工程教育评价指标体系构建及其应用——以 97 所“双一流”建设高校为样本[J]. 中国高教研究, 2021(2): 10-15.
ZHANG W, DENG Y X, XIN Y Y, et al. Construction and application of evaluation index system for higher engineering education from the perspective of multivariate classification: a case study of 97 “double first-class” construction universities [J]. China Higher Education Research, 2021(2): 10-15. (in Chinese)
- [11] CHEN C F. A study of college teachers' English teaching quality based on fuzzy neural network[J]. Computational Intelligence and Neuroscience, 2022, 2022: 8162048.
- [12] 牛富荣. 财政学翻转课堂教学质量评价体系的构建[J]. 山西财经大学学报, 2020, 42(S2): 125-128.
NIU F R. Construction of a quality evaluation system for flipped classroom teaching in financial science[J]. Journal of Shanxi University of Finance and Economics, 2020, 42(S2): 125-128. (in Chinese)
- [13] 张怡, 武小鹏. 基于 AHP-模糊矩阵的翻转课堂综合评价系统设计[J]. 现代远距离教育, 2018(5): 19-26.
ZHANG Y, WU X P. Design of comprehensive evaluation system of flip classroom based on AHP-fuzzy matrix[J]. Modern Distance Education, 2018(5): 19-26. (in Chinese)
- [14] 李爽, 刘紫荆, 郑勤华. 智能时代数据驱动的在线教学质量评价探究[J]. 电化教育研究, 2022, 43(8): 36-42, 76.
LI S, LIU Z J, ZHENG Q H. Exploration of data-driven online teaching quality evaluation in the intelligent era[J]. Research in Educational Technology, 2022, 43(8): 36-42, 76. (in Chinese)
- [15] 宋美喆. 地方应用型本科高校教学全过程质量评价框架的构建研究[J]. 科教导刊, 2019(16): 1-3.
SONG M Z. A Study on the construction of quality evaluation framework for the whole process of teaching in local applied colleges[J]. The Guide of Science & Education, 2019(16): 1-3. (in Chinese)
- [16] 任小宁, 郭晓黎, 杨伟东, 等. 高职中医专业实践教学教学质量评价体系的研究[J]. 卫生职业教育, 2019, 37(14): 64-65.
REN X N, GUO X L, YANG W D, et al. Research on the evaluation system of practice teaching quality of traditional Chinese medicine specialty in higher vocational colleges[J]. Health Vocational Education, 2019, 37(14): 64-65. (in Chinese)
- [17] YUAN T Q. Retracted algorithm of classroom teaching quality evaluation based on Markov chain[J]. Complexity, 2021(1): 1-12.
- [18] 余新桥. 基于表达式的 AHP 群决策在教学质量评价中的应用[J]. 信息技术与信息化, 2022(10): 111-114.
YU X Q. Application of AHP group decision-making based on expression in teaching quality evaluation[J]. Information Technology and Informatization, 2022(10): 111-114. (in Chinese)
- [19] 蔡红梅, 许晓东. 高校课堂教学质量评价指标体系的构建[J]. 高等工程教育研究, 2014(3): 177-180.
CAI H M, XU X D. Construction of quality evaluation system of classroom teaching in colleges and universities[J]. Research in Higher Education of Engineering, 2014(3): 177-180. (in Chinese)
- [20] 彭波, 王伟清, 张进良, 等. 人工智能视域下教育评价改革何以可能[J]. 当代教育论坛, 2021(6): 1-15.
PENG B, WANG W Q, ZHANG J L, et al. The education evaluation reform from the perspective of artificial intelligence [J]. Forum on Contemporary Education, 2021(6): 1-15. (in Chinese)
- [21] 吴立宝, 曹雅楠, 曹一鸣. 人工智能赋能课堂教学评价改革与技术实现的框架构建[J]. 中国电化教育, 2021(5): 94-101.
WU L B, CAO Y N, CAO Y M. Framework construction of artificial intelligence empowering classroom teaching evaluation reform and technological implementation [J].

- China Digital Education, 2021(5): 94-101. (in Chinese)
- [22] YANG W J, ZHOU Y J, YUAN S. Study of teaching assessment based on BP neural network [J]. Advanced Materials Research, 2012, 524: 3861-3865.
- [23] CHEN Y F, YU Z B, ZHAO W H. Based on optimization research on the evaluation system of English teaching quality based on GA-BPNN algorithm [J]. Computational Intelligence and Neuroscience, 2022, 22: 99-128.
- [24] JIANG L X, WANG X H. Optimization of online teaching quality evaluation model based on hierarchical PSO-BP neural network[J]. Complexity, 2020, 20: 664-683.
- [25] HOU J W. Online teaching quality evaluation model based on support vector machine and decision tree [J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(2): 2193-2203.
- [26] 苏超,王国中. 基于改进 OpenPose 的学生行为识别研究[J]. 计算机应用研究, 2021, 38(10): 3183-3188.
SUC, WANG G Z. Research on student behavior recognition based on improved OpenPose[J]. Application Research of Computers, 2021, 38(10): 3183-3188. (in Chinese)
- [27] 张鸿宇,刘威,许炜,等. 基于深度图像的多学习者姿态识别[J]. 计算机科学, 2015, 42(9): 299-302.
ZHANG H Y, LIU W, XU W, et al. Depth image based gesture recognition for multiple learners [J]. Computer Science, 2015, 42(9): 299-302. (in Chinese)
- [28] 林灿然,许伟亮,李逸. 基于多模态数据的课堂学生行为识别技术的探究[J]. 现代计算机, 2020(6): 69-75.
LIN C R, XU W L, LI Y. Research on action recognition technology of classroom students based on multi-modal data [J]. Modern Computer, 2020(6): 69-75. (in Chinese)
- [29] 周昊玮,刘勇,玄萍. 基于预训练和多模态融合的假新闻检测[J]. 计算机工程, 2024,50(1): 289-295.
ZHOU H W, LIU Y, XUAN P. Fake news detection based on pre-training and multi-modal fusion [J]. Computer Engineering, 2024,50(1): 289-295. (in Chinese)
- [30] LAN Z Z, BAO L, YU S I, et al. Multimedia classification and event detection using double fusion[J]. Multimedia Tools and Applications, 2014, 71(1): 333-347.
- [31] 王国中,郎亚坤,范涛. 机器助教打造课堂教育新模式[J]. 计算机教育, 2021(1): 142-145.
WANG G Z, LANG Y K, FAN T. Machine teaching assistants create a new model of classroom education [J]. Computer Education, 2021(1): 142-145. (in Chinese)
- [32] ZHOU L, WANG P Z, ZHANG C J, et al. Multi-mode fusion BP neural network model with vibration and acoustic emission signals for process pipeline crack location[J]. Ocean Engineering, 2022, 264: 112384.
- [33] XI T P. Research on the evaluation algorithm of English viewing, listening, and speaking teaching effect based on DA-BP neural network[J]. Mobile Information Systems, 2022, 2022: 4621405.
- [34] 张涛,刘晶,胡海波. IPSO-BP 神经网络模型的教学质量评价[J]. 实验室研究与探索, 2023, 42(6): 174-178, 193.
ZHANG T, LIU J, HU H B. Teaching quality evaluation based on IPSO-BP neural network model[J]. Research and Exploration in Laboratory, 2023, 42(6): 174-178, 193. (in Chinese)

编辑 索书志